

[illegible]

5' UTR: 1-134
Start: 135
Stop: 981
3' UTR: 984

Top 10 BLAST Hits

Score	E
301	6e-81
284	1e-75
283	1e-75
161	1e-38
161	1e-38
154	9e-37
154	1e-36
152	3e-36
151	8e-36
145	4e-34
141	9e-33
140	2e-32
140	2e-32

Station	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373</
---------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

FIGURE 1, page 2 of 3

[illegible]

Expression information from BLAST dbEST hits:

Expression information from PCR-based tissue screening panels:

FIGURE 1, page 3 of 3

```

1  MGAYLSQPNT VKCSGDGVGA PRLPLPYGFS AMQGWRVSME DAHNCIPELD
51 SETAMFSVYD GHGGEVALY CAKYLPDIK DQKAYKEGKL QKALEDAFLA
101 IDAKLTTEEV IKELAQIAGR PTEDEDEKEK VADEDDGDHF YKRKNLNPPE
151 EQMISALPDI KVLTLTDDHE FMVIACDGIW NVMSSEQEVVD FIQSKISQRD
201 ENGELRLLSS IVEELLQCL APDTSGDGTG CDNMTCIIC FKPRNTAELQ
251 PESGKRKLEE VLSTEGAEEN GNSDKKKKAK RD (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

233-236 NMTC

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 4

1	10-12	TVK
2	197-199	SQR
3	253-255	SGK
4	273-275	SDK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 10

1	38-41	SMED
2	57-60	SVYD
3	106-109	TTEE
4	122-125	TEDE
5	164-167	TLTD
6	184-187	SSQE
7	197-200	SQRD
8	210-213	SIVE
9	224-227	TSGD
10	229-232	TGCD

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 2

1	2-7	GAYLSQ
2	266-271	GAEENG

[5] PDOC00009 PS00009 AMIDATION
Amidation site

253-256 SGKR

[6] PDOC00792 PS01032 PP2C
Protein phosphatase 2C signature

55-63 MFSVYDGHG

BLAST Alignment to Top Hit:

```
>CRA|18000005093929 /dataset=nraa /length=546 /altid=gi|4505999
      /def=ref|NP_002698.1| protein phosphatase 1G (formerly
      2C), magnesium-dependent, gamma isoform; protein
      phosphatase 1G (formerly 2C);; protein phosphatase 2,
      catalytic subunit, gamma isoform [Homo sapiens]
      /org=Homo sapiens /taxon=9606
      Length = 546
```

Score = 301 bits (763), Expect = 6e-81
Identities = 146/146 (100%), Positives = 146/146 (100%)

```
Query: 137 GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 196
      GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI
Sbjct: 401 GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 460
```

```
Query: 197 SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNTAELQPESGKR 256
      SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNTAELQPESGKR
Sbjct: 461 SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNTAELQPESGKR 520
```

```
Query: 257 KLEEVLPSTEGAEENGNSDKKKKAKRD 282
      KLEEVLPSTEGAEENGNSDKKKKAKRD
Sbjct: 521 KLEEVLPSTEGAEENGNSDKKKKAKRD 546 (SEQ ID NO:4)
```

Score = 284 bits (718), Expect = 1e-75
Identities = 137/139 (98%), Positives = 138/139 (98%)

```
Query: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60
      MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD
Sbjct: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60
```

```
Query: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120
      GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR
Sbjct: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120
```

```
Query: 121 PTEDEDEKEKVADEDDGDH 139
      PTEDEDEKEKVADEDD D+
Sbjct: 121 PTEDEDEKEKVADEDDVDN 139 (SEQ ID NO:5)
```

```
>CRA|18000005157254 /dataset=nraa /length=542 /altid=gi|6679793
      /def=ref|NP_032040.1| fibroblast growth factor inducible
      13 [Mus musculus] /org=Mus musculus /taxon=10090
      Length = 542
```

Score = 284 bits (718), Expect = 1e-75
Identities = 139/146 (95%), Positives = 141/146 (96%), Gaps = 1/146 (0%)

```
Query: 137 GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 196
      GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI
Sbjct: 398 GDHFKYKRNKNLPPEEQMISALPDIKVLTLTDDHEFMVIACDGIWNVMSQEVVDFIQSKI 457
```

```
Query: 197 SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNTAELQPESGKR 256
      SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNT ELQ ESGKR
Sbjct: 458 SQRDENGELRLLSSIVEELLDDQCLAPDTSGDGTGCDNMTCTIIICFKPRNTVELQAESGKR 517
```

Query: 257 KLEEVNSTEGAEENGNSDKKKKAKRD 282
 KLEE LSTEGAE+ GNSD KKKAKRD
 Sbjct: 518 KLEEALSTEGAEDTGNSD-KKKAKRD 542 (SEQ ID NO:6)

Score = 279 bits (706), Expect = 3e-74
 Identities = 133/139 (95%), Positives = 138/139 (98%)

Query: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDSETAMFSVYD 60
 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELD+ETAMFSVYD
 Sbjct: 1 MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPELDNETAMFSVYD 60

Query: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALEDAFLAIDAKLTTEEVIKELAQIAGR 120
 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKAL+DAFLAIDAKLTTEEVIKELAQIAGR
 Sbjct: 61 GHGGEVALYCAKYLPDIIKDQKAYKEGKLQKALQDAFLAIDAKLTTEEVIKELAQIAGR 120

Query: 121 PTEDEDEKEKVADEDDGDH 139
 PTEDED+K+KVADEDD D+
 Sbjct: 121 PTEDEDDKDKVADEDDVDN 139 (SEQ ID NO:7)

Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF00481	Protein phosphatase 2C	176.2	4.2e-49	2
PF01722	BolA-like protein	3.7	9.5	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00481	1/2	25	102 ..	1	95 [. .]	74.1	4.6e-20
PF01722	1/1	99	110 ..	76	87 .]	3.7	9.5
PF00481	2/2	137	228 ..	197	301 .]	102.2	4.5e-28

```

1  AAAGAATCTT TTTTTTTTTT TTGAGACGGA GTTGCTCTGT CACCCAGGGT
51  GGAGTGCAGT GGCGCCATCT TGGTTCAGTG CAACCTCCGC CTCTGGGGTT
101 CAAGCTATTC GCCTGCCTTA GCCTCCCAAG TAGCTGGGAT TACAGGAGCG
151 CACCACTACG CCTGGCTAAT TTTTGTATTT TTAGTAGAGA CGGGTTTCAC
201 ATGTGGGCA GGCTGGTCTC GAACCTCTGG CCTCAAGTGA TCCACCACCC
251 CCCTTGGGC TCCCAAAGTG CTGGGATTAC AAGTGTGAGC CACTGTGCCC
301 GGTGAAAAG AATCAATTTT GTCATAGTTT GGAGAATTC TCCTTTTCTC
351 TCCATCCCTT GAATGCAATT TATTACCAA TCTGTCTTAT TTGTTATTGT
401 CTAATTTGTC CTTTCATCTG GATTCCCATT GCCACCTGC GTGGTACCAC
451 CTTACTCCCA GCTCTTCTCA TCTCCTGCTT AGAGTAAGAG CTCTCTAACT
501 AGTAGCAGTG CCCAGGCCA GCGCGGTGG CTCACGCTG TAATCCCAGC
551 ACTTTGGGAG GCTGAGGCGG GTAGATCACG AGGTGAGGAG TTCGAGACCA
601 GCCTGGCCAA CATGGTGAAG CCCCGTCTCT ACTAAAAATA CAAAAATTAG
651 CCAGGCGTGG AATCCTAGCT ACTCGGGAGG CTGAGGCAGG AGAATCGCTT
701 GAACCTGGGA GCGGAGGTT GCAGTGAGCC AAGATAGCGC CACTGCACTG
751 CAGCCTGGGC AACAGAGCG AAACCTGTGC TAAAAAATA AACAAATAGT
801 GCGGGGGTGC GGTGGCTCAC GCCTGTAATC CCAACACTTT GGGAGGCCGA
851 GCGGGGCGGA TCACGAGGTC AGGAGATGGA GACCATCTG GCTAACACGG
901 TGAACCTCG TCTCTACTAA AAATACAAA AATTAGCCAG GCGTGGTGGC
951 GGGCGCTGT AGTCCCAGCT ACTCGGGAGG CTGAGGCAGG AGAATGGCGT
1001 GAACCCGGGA GCGGAGCTT GTAGCCTGGG CGACAGAGCG AGACTCTGTC
1051 TCAAAAACAA AAAAAAACA ACAACACAAC AGTGCCCCAG ACTCTCTCCC
1101 TCCAATGTAC ACTGCATACA AAGACTAGAC AAACGATGCC AAAGGTTACA
1151 CCAGGACGAC AATGAAGTCC AAGTCACTCA TCTTGGCGCT TCTTTTATC
1201 AAGCTAGCTA GTTAATATTA CCACCTACAA TCATTCTCCG AGTCCAGCCA
1251 GCTTCCTCAG AAATCCTCCC CAAATGCAGT TCACATTCTT CCCTCACTCC
1301 TCCAAACCTT ATAACATTAT TCCCTTTTCC TGTGGCATAA TGCAATCCAG
1351 AGGCATCCTC TCTGAGAGCC TACATCCTTT CAATCCTCCA AGAAACAGCT
1401 CCTCCTCCCA TCCTTGAGCT TTTCCCAACC AGAATAGGCT GTACCAAACA
1451 TTTCTACATT GTATTAACTC GAATTATATC ACAGATCGAG GTCTTCGCTT
1501 TCAGAAAAGA CTCACATTCT TCTCATATAG GCCTCGCATG GCCTGGCCAA
1551 GTACTTCCTA AGCTCGGAAC AAATACTGGT CAACTTTAAT TGAACCAAT
1601 CGGGCGGGGT TTGCGGAGTC TGAGAGTGCA ACGTTGGGGA GAGGGGGATG
1651 AAAAAGCTGG AGGACGGGCT GAAAGCGTCG AGTCCGACAC AAAAGAGGCG
1701 TCAGACAAAA CGCCAAGAGG CTGGGGACTG GGAACGAAGG AAGAGGTTCT
1751 GCCAGAGGCG ACCTGCCACC TGCGCGAGGA AGCGGAGTAG GACGGCGGCC
1801 GTTGGTGGGC GTGGTCGCGC TAGTCTCGCG GGAGCGGCCG TTGGGCGGGC
1851 CGTTGTCCCC TGCGGGCGGG GCGAGTTGCT AAGGAAATGA CTGCCCGCAG
1901 CGCCTGGCCC CGCCGCGCAG GCCGGGCGGG GTCTGGAGCG GCGCCGTTTC
1951 CGCTTCCGCT CCCTCACAGC TCCCGTCCCG TTACCGCCTC CTGGCCGGCC
2001 TCGCGCCTTT CACCGGCACC TTGCGTCGGT CGCGCCGCGG GCCTTGCTCC
2051 TGCCGCGCGC ACCCCGCGGG CTTCGGCTCC GGCACGGGTC GCGCCAGCT
2101 TTCTTGACAC TGAGCCGCGG GCCAGCCGCC GCCATGGGTG CCTACCTCTC
2151 CCAGCCCAAC ACGGTGAAGT GCTCCGGGGA CGGGGTCGGC CCCCAGCGCC
2201 TGCCGCTGCC CTACGGCTTC TCCGCCATGC AAGGCTGGCG CGTCTCCATG
2251 GAGGTGAGGA GGCAGGGGCC CATAGGCTGG CCGCTGCGGG GCGGGAATCT
2301 GACGGAGAAA GAGAGCGGGG GATGGGTCC TCCCCTGGGA AGGGTCCCAA
2351 TTGGGAGCCT GCGGCCGAG CGGCCGTTTG CGGGGCGACA GAGACCGCGG
2401 GGTGAGGGCC GAGAGGGAGC TCCCGATGCT TGGGGACCGT GTGCCGGTGG
2451 CCAGTGGCGG CGAGGGCTCG GCCATGTGGG AAGAGGCACC TTCCGCCAC
2501 TGACCGCCCT CTCCCAGAGC TTTGGCGCCA TCCTCTCGTG CCAACCTAGC
2551 CCTCCAGGCT CATCAACGGT GTGAGGTTTA GTGTGGGAGT AAAGACGCAA
2601 ATAGGGGCTA TTTATTCATT TGTTTTTCAA GGGAGAGGTG AATAGATGTG
2651 AATAACTTTT TAAATTTTAA TATTTAAAT ATCTGATGTG GGAAGCCTCT
2701 TTTGGCTAGG AGTTTGACAG TGAAAGGAAC CCCGGGCGAG GTCTGTTTCA
2751 CATTTTGGTT GCCTGGCCTT GGGCTCTTGC ATGTAAATTT CAGAGGCTGG
2801 ACCCGACCTC CAGGAGTTGT CACTCATTG CACTCTTTC AGGCCTTTT
2851 ACTAACTTCG GAAACTTGA ATTATGTCAG TCCCTAGGTT TTCCTTTTAA
2901 TTATGCTTCG TTTTCTTCT CTCTTAGGAT TTCTCTAAAA CTTAATCAGT
2951 AATTCTCATT TGCCTGTAAT TGTAAGTTAT AGTTTCTTTG ATTGTAAGTC
3001 ATTTTCATCA TTTTCTTCAT CACAACCTAC CTACAAAGGG CTTTCTAGA
3051 AAATTTTACT CTGGACAAAA GGGGAAAAGA AAAATATTGG GGGGAAAGTA
3101 GTAGTATTAG GTAAACTTG ATGTGAAACT ACAAAGAGA AGAGGAAAA

```

FIGURE 3, page 1 of 17

```

3151 CTGCGGTAGG GAGGAAAGGG AGGAAGACGG GTTAACCGTG GCTTTGTGAA
3201 GAGCATTCTG AAGTCTAGGC AAAAGGGCCA GGGAAATACT CTGTCTGGTA
3251 TTGAGGGTTT CTCCACCTAC CGGGTGGGCT TCAGGTAACA GCGAAATACT
3301 GTCTCCCTTG GGAATTGTTT CAGATCCCTC GCTCCTCTG TGGTTAGCTC
3351 TGGAATGCCA GTATGAACCT CAATGTTTTG TTTTCCGATT CAAATTTTAT
3401 ATTCATAACT GACCTTAATA ACAATTTTAC AATTAGGTAT AAAATTTTCTCAG
3451 GATCCTAGTG TATCCTATAG TTCATCTCAT CTGCTTTGGC TCCCTTTTTT
3501 TTTTTTTGAG ACAGAATTTC GCTCTGTGTG CCCAGGCTGG AATGCAATGG
3551 CGCGATCTCG GCTCACCACA ACCTCCGCCT CCCAGGTTGG AGCAATTCTC
3601 CTGCCCTCAG CTCCATAGTA GTTGGGATCA CAGGCATGTG CCACCACGCG
3651 TGGCTTATTT TGTATTTTGA GTAGAGACAG GGTTTCTTCA TGTGGGTGAG
3701 GCTGGTGTG AACTCCTGAC TTAGGTGATC TGCCTGCCTC GGTCTCCCAA
3751 AGTGCTACCA CCACNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
3951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
6001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
6051 NNNNNNNNNN NNNNNNNNNN NNNNNNTGGA TTACAGACAT GAGCCACTGC
6101 GCCCAGCCTT ATTTAGAAAT TTCTTTAGTG AAAGATGATA AATTTTCAGT
6151 TTTTCATTAT CTGAACATGT TTTTATCTAG CCTTTGTCT GAAAAGATGC
6201 TTGGACTCAG TACCCAGTTC TAGATTGACA GTTAATTTT CTTAATTTGT
6251 AAATGTTGTT TCATTGATTG ACTTCCATTG TTGTTTCGAA AAATTTATCA

```

FIGURE 3, page 2 of 17

6301 TCAGCCATTT CTGACTTTTG ATCTGTGTTT TCTCTTTGGT TTCTCTTTTT
6351 TTTTTCCTTT TTTTTCCTTT GAGACGGAGT GTCGCTCTGT TGCCAGGGCT
6401 GGAGTGCAGT GGCATGATCT TGGCTCACTG CAACCTCTGC CTCTGAGTT
6451 CAAGCGATTC TCCTGCCTCA GCTTCCCGAG TAGCTGGGAT TACAGGCGCC
6501 TGCCACCATG CCGGCTAAT TTTTGTATT TTCAATAGAG ACAGGGTTTC
6551 ACTATGTTGG CCGGGTTGGT CTCCAACCTC TGACCTCTTA ATCCGCCCGC
6601 CTCGGCCTCC CAAAGTGCTG GCATTACAGG CGTGAGCCAC CATGCCTGGC
6651 CCATCAGTTG ATGTAGTCTT AAGGGGACAA GAGTACATTT AATATTTGGT
6701 TGTGAGGTTT TCTGGAAGTG ACAAACCTGC TTTCTATGGA GAGTTAGGAA
6751 TTTTTCCTTT TTTTGAAACG GAGTCTCGCA TTGTCACCCG GGCTGGAGTG
6801 TAGTGGCTTG ATCTCGGCTC ACTGAACTT CCGCCTCTTG GGTCAAGTG
6851 ATTCTCCTGT CTCAGCCTTC CAAATAGCTG GGATTACAGG AGTCTGCCAC
6901 CAGGCCAGCT AATTTTCTTT TGTATTTTAA GTGGAGACAG GATTTCACTA
6951 TGTGGCCAG GCTGGTCTCA AGACTCCTGA CGTTGTGATC CACCTGCCTC
7001 GGCTCCCAA AGTGCTGGGA TTACAGATGT GAGCTACCGT GCCCGGCCAG
7051 GAATTTTCTG TGCTATAAAT CATATTTTCC TTTTATTTAA GGCAGTGTCA
7101 ATATCTATAG TATAATTTTG AGGAGGCTGG CTATTTATTG CTGTGTAGAA
7151 GCTGGCTTAT TAGTGGTCAA GGGGTCTCT AGAATTGACT ATAAAGATAG
7201 TATTGAGCAG AAAATTTCTT AAATACCTGC ATATTAGTTT CAGTCATTAA
7251 ATTAATGGAA AAAATATAAA AAGAAATATC ACAAGTATGC TATGGGGTTC
7301 TACCTTAGGG CATTGAAGGT TGAAAAACAT TTATTTTCTT ATCTTCAGAT
7351 TAGCATCTCA TATCAATAGC CAATAGCTTA AAGCGCTTTT TACTTACTAA
7401 ACCAGGTCAG AATTTCTCTC TCTCTCTCT TTTTTCCTTT AGACAGAGTC
7451 TCTGTACACC AGTCTAGAGT GCAGTGGTGT GATCTCGGCT CACTGGCCTT
7501 TCTGTTTAAG TGATTCTCCG GCCTCAGCCT CCGAGTAGC TGGGATTACA
7551 GGCATGTGCC ACGACACTTG GCTTATTTT TGTATTTTAA GTAGAGATGG
7601 GGTTCCTCTG TGTGGTCTAG GCAGGTCTTG AACTCCTGAC CTCAGGTGAT
7651 CTGCGCGCCT CCGCTCCCA AAGTGCTGGG ATTACAGACG TGAGCTACTG
7701 CGCCTGGCCA GAATTTCTTT GTCTAGAATG TGGTTAGCAA CTTTATAAAA
7751 AACGCATTAT TTGCATTTGA TTAGCATGCA GTACCCATTC ACAGTTCAAA
7801 GCTAGTATAG AATTATATCA CATGTATGCC CATGAGCATG GAGAACTAT
7851 TTTCTTTTAA TTTTCTTAAG TTGGAGTTT GCTCTTGTG CCCAGGCTGG
7901 AGTGCAATGG TGCCATCTCG GCTCACTGCA GCTTCTGCCT CCTGGGTTCT
7951 AGCAATTTGC CTACCCAGC CTCCCAAGTA GCTGGGATTA CAGGCACTCG
8001 CCACCATGCC CAGCTAACTT TTTTGTATTT TTAGTAGAGA AGGGGTTTTT
8051 CCATGCTGGC GAGTCTGGTC TTGAACCTCT GGCTCAAGT GATCTGCCCC
8101 CCTCAGCCTC CCAAAGTGCT GGAATTACAG GCATGAGCCA CTGTGCCCCG
8151 CCTTTTATTT TTTTAAATTA TTTATGTATT TATTTTGAGA CAGGATCTCA
8201 CTCTTGCCCA TGCTTGAGTG GTAAGGAGTA TGGGATTCTT TGTGCCGGTT
8251 CCTTCACATC CTCACTACAC TTATCTGCCT TTACAGTGGC TCGATCATAG
8301 TTCCTGTCAT AGCCTTCTGG GCTCAAGGGG TCTTCCAGCC TCAGCCTAAT
8351 ATAGGCACAT GCCACCATGA CTGGCTAATT TTTTTCCTTT AAGTTTTCCT
8401 TTGTAGAGAT AGGGCCTTGC AGTGTGCCCC AGGCTGAGGA ATTTTATTTA
8451 TGTATTTTAT ATTTATTTAT TTATTTATTT ATTTATTTAT TTATTTATTT
8501 ATTGAGATGG AGTCTTACAC TGTCACCCAG GCTGGAGTGC AGTTGCGCGG
8551 TCTCAGCTAC TGCAAGCTCC GCCTCCCGGG TTCATGCCAT TCTCCTGCCT
8601 CAGCCTCCCG CGTAGCTGGA CTACAGGCGC CCGCCACCAC GCCTGGCTAA
8651 TTTTGTGTGT GTGTTTCTAG TAGAGGCGGG GTTTCACCAT GTTAGCCAGG
8701 ATGGTCTCGA TTCTGACCT CGTGATCCAC CCACATCGGC CTCCCAAAGT
8751 GCTGGGATTA CAGGTGTGAG CCACCATGCC TGGTCTAGAA ATTATTTTAT
8801 ATTTTATACC ATTTGCTTAT AAGTTCTCAA GCAACTGGAA AATACAATCA
8851 GAACGTATTC CTCAAGATTT CAAGGATATT TTACACAAAG TTCTATTGTC
8901 TGATTCCTTA GCAGTTGTTA CTAAGTTTTC CCTAACCTCT AATCTTCTAT
8951 TGGGTTATTA GTCTTAGAAT TGAATTTTGA GAGGTAAGGG CTTGAATTTG
9001 AACATAGAAA TTTATACAGG TCTGATCAGT AGTTCTTGAC ATTGTATTAT
9051 CTGGAAACAA ATCTTTAGAA CTGAGCTTAA GATGTTTAAAT GACATTTTGT
9101 AGACAGAGTA TGATTTCACT GTAGTTGTTT TTGTTTCTTT CTAGATCTAG
9151 TTCAGAGATG AAGTATATCA ACTTTTTTTT CCTTTTGTAC CCAATGCTAG
9201 CAGAAAAACA ACACCTTTTA ATCATATTTA GTATTTGAAA ATGTGTATAC
9251 AGGTTCCCTT TTATTTTATT TATTTCTTAC AGGTTCCCTT TTAATCAGCT
9301 TTATTGAGAT AGAGTTCATA TACTGTATGG TTCATACCAC ATATGGTTCA
9351 TATACCATAC AGTATATGAA CTCACTTTAA GAGTATAAAT CAGTGGGTTT
9401 AAGGATATAA TTCATTCAAT TTAAGGTAT AATTCAGTGG CTTTATAGTAT

FIGURE 3, page 3 of 17

```

9451 ATTTTCTTTT TTTTCTTCTT TTTTCTTTT TTCGAGACAG GGTCTTGCTC
9501 TGTTACCCAG GCTGGAGTGC AGTGACGCAG CCTCAGCTCA CTGCAACCTC
9551 CACCTCCCAG GTTCAAGCGA ATTCTCTTGC CTCAGCATCC TGAGTAGCTG
9601 GAACTACAGG CTCACGCCAC CACACCCAGC TAATTTTAT ATTTTCAGTA
9651 GAAACAGGGT TTCACCATGT TGCCAGGCT GGTCTCGAAC TCCTGGCCTC
9701 AAGTGATCTG CCTGCCTCAG CCTCCCAAAG TGCTGGGATT ACAGGCATGA
9751 GCCACCGTGC CTGCCTGTTT TGTAAGTAT TCAAACAGAG TTGTACAAC
9801 GTCACCACAA TCAGTTTTAG AACCCCAAAA AGAAACCCTG TACTCTTTAC
9851 CAGTCACTCC CTATCTTCCG TCCACTAACC CCTGGCATCC ACTAATTTAC
9901 ATGACCTCTA TGAATTTGCC TATTCTGAAC ATTTTATAAA TGGAATTCTA
9951 AATACACTAC CTTTATATATC TGGATGCTTT TACTAAGCAT GTGTATTTT
10001 GAAATTGACT TTAAGCTTG TTGGCCCTG GAAGAGTAAA TTAATCTCCA
10051 CCCCAGTAT TCCCTCTACC CCTCAGCTTT GCCTGTAAGT TTCTTTTAA
10101 AAAATCACA CATACATTGT TGTAAGTAT TTAAGAATAA GTATTTTGCT
10151 GACCCAAGGT TCTTTTGCTT CTTTCTAGAT CAGTGCCTTG CAGGTTTTAT
10201 TTTACAGAGC TTAATAGAAT CAGAAATCTC TTTAAACTC CAGTCTCATA
10251 TCCAGTTATC ACTCACCATC TCTGTGTTT CAGCAATAGC CAGGCCTGGC
10301 CCAGAGGGAC TTGATCTCCA CTTTGGTTT TTAGACTTTT CTGTGGCTTT
10351 TACCACCTGC TGTGTATCCT TGACCTCATA CTGCTGGACT CCTTTGGATG
10401 GATACCAGCA GGATGGTTCA GGCTCCAGTG GGCACCTTTT AAAATTCTCT
10451 CTTTCTGTTT AGATAGACAG AGCTCAGGCA GATCACCAGT TCTGTTGCC
10501 GTGTAACCAG GAAAGAGATG CTAGTTTCTT TTAGGCACT CCCATTTGTT
10551 TCTGTTGGAA CCTTCTCAC TTAGTTGATG GAAGGGAAGC AAAAGACCCA
10601 GAACCTCATT TCAAATTAAT GACTTAACAA TTCTTGAATT TTCTCTTATC
10651 TCCTAGTTAA CTCTTTTCT TATCTCCAGG AGGTCACTTT TAATTATTGT
10701 TGTTTATTCA TTTATTTTCA TGGAGACAGG GTCTTACTGT GTTGTCCAGG
10751 CTGGCCTTGA ACCCCTGGCC TCAAGGAATC TTTTCACTC AGCCTCCCAA
10801 AGTGCTGGAA TTACAGGCAT GAGTCAACCAC ACCCAGCCTG ATATTTTTCA
10851 GTTGATGTAT CATAGTTGTG CCTAAGCATA ATTTTAAAT TTTAATTTTT
10901 TATTTTGGG GACAGGGTCT CCCTCTGTCG CCCAGGCTGG AGTGCAGTGA
10951 TCGATCTCA GCTCACTGCA ACCTCCACCT TCTGGGTAA AGCGATTCTC
11001 CTGCCTCAGC CTCCCGAGTA GCTGGGACTA CAGGTACCCA CCATCACACC
11051 CGGCTAAATT TTTTGTGTGT ATTTTAGTAG AGACGAGATT TCGCTGTGT
11101 GCCCAGGGCG GTCTCGAAT CCTGAGCTCA GGCAATCCGC CTGCCTCGGC
11151 CTCCCAAAGT GCTGGGATTA CAGGCATGAA CCACCACGCC CGGCCAAGCG
11201 TAATATTTT AAGGGTCATC AATGTTGTGT CATGAATCAA TCAGTGTTC
11251 GTTCTTTT ATGGTTGAAT AATATTCCAT GGTATGGATT TGTACATTT
11301 TGTATATCCA TTCAATAGTT GATAGACATT TTGGATTCC ACTTTTTTTT
11351 TTTTTTTTT GCTATTATAA ATAGTGATAC TATGTACAAA TTTTGTGTG
11401 GAAATATGTC CTCATATCTC TTGGTTATAT ACCAAAGAGT GGAAGTGCTG
11451 GGTCAATGG TAACTACGTG TTTAACATTT TGAGAACTG CTAAGTGT
11501 TTCCAAAGTT GCTGTACCGT CGTACATTCC TGCCAGCAAT ATATGAGGAT
11551 GCCAGTTTCT TCACATGTTT ACTACACTTA TCCACCTTTT TTATAATAAC
11601 TAATGGTGGG TGTGAGATGG TATCTCATTG TAGTTTTGAT TTGTATTTCT
11651 CTGATGGCTA AATGGCTAAT GATGTTTGAA CTTTTTGTG GAGACAGAAT
11701 CTCACTCTGT CCAGATTCAA GCGATTCTCC TGCCCTCAGC TCCCTAGCAG
11751 CTGGGATTAC AGGCACATGC CACCACACC AGCTAATTTT TTGTATTTCT
11801 AGTAGAGACA GGGCATTACC ATGTTGTTCA GGCTGGTGTG GAACTCCTGA
11851 CCTCAAAGGA TCCGCTCCC TGGGCTCCC AAAGTGCTGG ATTACAGGCT
11901 AGAGCCACCA TGCCAGGCTT TATGTTTGAA CATCTTTTAT GTGCTTATTG
11951 GACATTTGTG TATCTTCTTT GGAGAAATGT CTGTTCAAAG TCTTTGTCCA
12001 TTTTAAATTG GATTGTCTTT TTGTCTTTTG ATGTGTAAGA GTTCTTTATG
12051 TGTTTTGGAT ACAAGTTTGT TAGATATATG ATTTGCAAAT CTTTCTCCA
12101 ATTTTGTGG ACTTTTGCTT TCTTTTTTTG TTTTGTTTT GTTGTGTTG
12151 TTGTTGTTGT TGTTTTGGTC GGGGACAGT CTGCTCTGA CCACCCAGGC
12201 TGAATGGAG TGGCGCATC TTGGTTCACT GCAACCTCTG CCTCCTGAGT
12251 TCAAGCTATC CTGCTTCAGC CTCCGAGTA GCTGGGACCC AGGTGTGTG
12301 CACCACTCCC AGCTAATTTT TTATTTTAG TAGAGACCGG GTTTCACCAT
12351 GTTGGCCAGG CTGGTCTTGA ACTACTGACC TCAGGTGATC TGCTGCCTC
12401 AGCCTCCCAA AGTGCTGGGA TTACAGTCAT GAGCCACTAC ACCCTGATTC
12451 TTTTGTCTG CTTTCTTTCT TTTTTTTCT TTTTTTTTT GAGACGGAGT
12501 CTCGCTCTGT TGCCAGGCTG GAGTGCAGTG GCATGATCTC GGCTCACTGC
12551 AACCTCTGCC TCCCGGTTT AAGCCATTCT CCTGCCTCAG CCTCCCGAGT

```

FIGURE 3, page 4 of 17

12601 AGCTAGGACT ATAGGCACAT GCCACCATGC GCAGCTAATT TTTGTATTTT
12651 TAGTAGAGAC GGGGTTTCAC CATGTTGGCC AGGATAGTCT CGATCTCTTG
12701 ACCTCGTGAT CCGTCCGCCT GGGCCTCCCA AAGTGCTAGG ATTACAGACG
12751 TGAGCCACCA CACTCAGCCT CTTTTTGCTT TCTTGATGGT GTCTTTTGAA
12801 AAAAAAGTTT TTTACTTTGA TAAAGTCCAA TTTGTCTATT TTGTTTGT
12851 GTTTTTGTGA AGAAGCTTTG CCTAACCCAA AGTCACGAGA ATTTTCTCTT
12901 AGGTTTCTCT CTAAGAGTTT TATAGTTTAA GCTGTTTCTG TGATCCATTT
12951 TGAGTGAATT TTTGTGAATG GTATGAGGGA GTGATCCAAC TTCATTCTTT
13001 TGTGTGTGGA TATCAAGTTG TCCCAGCACT ATTTGTTTAA ACCACTGTTC
13051 TTTTCCCCCA TTGAATTATC TTGGCATCAT TGTCAGAGAT AAATTGACCG
13101 TAAATGTGAG GGTTTTATTT CTGAACCTCTC AAGTCCATTT CATTTGGTCTA
13151 CATGTCCCTA TGCCAGTAAT AACTATCTCT GGTACTGTGA GCTTTTGTAGT
13201 ACGTTTGTGA ATGTTTTTAA AATTTGTTTT TCATCTAAAT TTTAGGATTA
13251 ATTTGTCAAT TTCTGCACAA AAGGCACCTG GGTTTCTATA GGGGTATATGC
13301 AGAATCTGTA GATCAACTGG GGGAGTATTA CAGGCATGAG CCACCGTGCC
13351 TGGCTGACTG AGTTTTTCAT AGATGTACTC TATCAGGTTT AGGAAGTTCC
13401 CTTTTATTCC TAGGTTGTTG AGTCTATTTT ATATTACTTT TTTAGAGACA
13451 GTCTTGCTCT GTCCCTCAGG CTGTAGCACA GTGGCTCAAT CATAGCTCAC
13501 TGCAGCCTTG AACTCCTAGG TTCAAGAGAT CCGCCTGCCT CAGCCTTCTT
13551 AGTAGCTGGG ATTACATGCA TGCACCACCA TACTGGGCTA ATTTTTTAAA
13601 ATTTTTTATA GAGACAGGGT CTTATTACTA TGTGCCCAG ACTGGCATTG
13651 AGTCTTTTAA TCATTAATGA GCACTGAATT TTGTCAAGTG CCTTTATAAT
13701 ACCTATTGTG ATGATCATAG GGTTTGTGTC TTAGTCTAC CGATACGCTA
13751 TATTGCATTA AGTGATTTTT TTGAATGTTA AACCAACCTT GCATTTTTTT
13801 GGTGTATAAG TCTTATTGTA TCAATGTGTA TTATCCTTTT ATATGGTGCT
13851 GGATTTAGTT TGCTACTATT TTGTTGAGGA TTTTGTGTCT TATATTCATA
13901 AGAGATATTG GTCTGTAGTT TCTTGTGATG TCTTGTCTG GTTTTAGAAT
13951 CAGGGTAATG CTGGCCTCAT AGAATGAATT GGGAGTGTG GTCTTTCTA
14001 TGTGATGGGA GAGTTTGTGA ATCATTGGTA TTAATTTTTC TGTAATGTG
14051 TGGTAGAATT CACAAATAAA GGCATCTGAG CCTGGGCTCT TCTTTGTGGG
14101 AAGTTTTTGG CTTTTTTTTT CTTTAAAAAT TTTCATTGTG GCTGGGCATG
14151 GTGGCTCAGC CCTGTAATCC CAGCACTTTG GGGGGCCAAG ACGGGTGAAT
14201 CACCTGAGGN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNGAG
14501 TGCAATGATT GCAGTCTTGG CTCACTGCAA CCTCTGCCAC CTGGGCTCAA
14551 GAGATTCTCC TGCCCTCAGCC TCCTGAGTAG CTGGGATTAC AGGCATGCGC
14601 CACCATGCCC AGTTAATTTT TGTATTTTAA GTAGAGACGG GGATTCTCCA
14651 TGTTGACCAG GCTGGTCTCT AACTCCTCAC CTCAAGTGAT CCGCCCGCCT
14701 CGGCCTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACCA CGCCCGGCCT
14751 TAAAAATTTT TTTAATGTAC AGTTGAGTAG TATTTAATAC ATTCACATTG
14801 TTGTGTACCC AGTTTCCAGA ACTCTTCATC CTACAGAACT GAAACTCCAT
14851 ACCCATTTAA TGAGTCCCCA TTCTCTTTCC CCCAGTCTAT GGCAACAGC
14901 ATTCTATTTT CAGTCTCTAT GAATTTGATT AGTTTAGATA CTTCACTAG
14951 TAAGTGAAT CATATGGTAT TTGTCTTTTA GTGACTGCCT AATTAAAAA
15001 AAATTTTTTT GAGACGGAGT CCTGCTCTGT CGCCAGGCT GGAGTGCAGT
15051 GGCACCATCT CTGCTCACTG CAACCTCCAC CTCCCAGGT CAAGTGATTG
15101 TCCTGCCTCG GCCTCCACG TAGCTGGGAT TACAGGTGCT CGCCACAACA
15151 CCCGGCTAAT TTTTGTATTT TTAGGTAGAG ACTGGGTTTC ACCATGTTGG
15201 CCAGGCTGGT CTCGAATCC TGACCTCAA TTATCCACCT GCCTTGGCCT
15251 CCCAAAGTGT TGGGATTACA GCGTGAGCC ACTGTGCCA GCCTCCATGT
15301 TGTTTTTCAC AACACCTGTA TCATTTACAT TTCCACCAAC AGTACACAAG
15351 AATTTTCACT TCTCCACATC CTTGCTAGCA GTTGTATTA TCTGTTTTTT
15401 TTTAATGGTT TCTTTTTTCC TTTTCTTTT TTTTTTTTTT TGAGACGGTC
15451 TTATCTCTGC TCATGCTGGA GTGCACTGGT GCAATGTGAT AGCTCACTGC
15501 AGCCTCAACC TCTGGGCTCA AGCAGTCCTG CCACCTCAGC CTCCACATAG
15551 GTGGGACTGC AGGTGTGCAC CACCACTGTG GGCTAATTTA AAAAATTTT
15601 TCGTAGAGAC AGAGTCTCAC AGTGTACCC AGGCTGGTCT TGAAGTCTG
15651 AGCACAGTG ATCCTCCAC CTCAGCCTCC CAAAATAATG AGATTAGAGA
15701 CATGAGCCAA CATGCCCAAC CAGTTTTGTT TGTGTTTTT GTTTTGTGTT

FIGURE 3, page 5 of 17

```

15751 TTTGAGACAG AGTCTCACTC TATTGCCAG GCTGGAGTGC AACGGCATGA
15801 TCTTAGCTCA CTGCAACCTC CGCCTCCAG GTTCAAGTGA TTCTCAGGCC
15851 CCAGCCTCCT GAGAAGCTGG GATTACAGGT GTACCACCAC ACCCAGTTAT
15901 TTTTGTATTT TTAGTAGACA TGGGGTTTTG CCATGTTGGC CAGGCTGGTC
15951 CCGAACTCCT GACCTCAAGT GATCTGCTCC CCTCAGCCTC CCAAGGTGCT
16001 AGAATTAAGT TTTTCTTTCT TTCTTTCTTT CTTTTTTTTT TTTTTTTTGA
16051 GACAGAGTCT CACTCTGTCA CCCAGGCAGG AGTGCAATGG CACGGTCTTG
16101 GCTCATGTGA ACCTCTGCCT CCCAGATTCA AGTAGTGATT CTCCTGTCTC
16151 AGCCTCCCAA GTAGCTGGGA TTACAGGCAT GCACCACCAC GCCCAGCTAA
16201 TTTTTTGTAT TTTTAGTAGA AACGGGGTTT CACCATGTTG GTCAGGCTGA
16251 TCTCAAATC CTGACCCCAA GTGATCCACC CGCCTTGGCC TCCCAAAGTG
16301 TTGGGATTAC AGGCGTGAGC CACTGTGCCT GGTTTTATTT TTATTATTAT
16351 TATTTTAAAT AGTTCCTATT CTAATGGGTA TGAGGTAGTG AGGTGGGTGG
16401 TTGTGGTGTT TTTATGAATG TTTAATGGGA AATGGGTGGC CATTGTGTGC
16451 AGGAAAAACC TCCTAAATTG TGTCAAACTC CTGGAAAAATG AAATATCATT
16501 CCAGTTGCAA GAATATCTTT TTTTTTTTTT TTTTTTTTTT AGACAGAGTC
16551 TCACCTCTGT ACCAGGCGGG AGTGCAGTGG CAGGATCTCG GCTCACTGCA
16601 ACCTCCGCCT CCTGGTCCGC CTCCCGGTT CAAGTGATTC TTCTGCCTTA
16651 GCCTCCCAAG TAGCTGGGAC TACAGGCGCG TGCCACCCTT CCTGGCTAAT
16701 TTTTTTGTAT TTTTAGTAGA GATGGGGTTT CACCATGTTG GCCAGGATGG
16751 TCTCAATCTC TTGACCTAGA GATCCGCCTG CCTCGGCCTC CCAAAGTGGT
16801 GGGATTACAG GGGGGTCACC GTGCCAGCC ACAAGAAGAT CTTGAGCATG
16851 TGAATGATCA GAAATGATTT AGCCTATGTA GGCCTAGGC CAGGTAGTGA
16901 AATTACAGGA AAATAATTCA GATGCTTCTG AGCTATCACT TATGAATAA
16951 GAAACAGCTT AAAGCCATTA TAGTGTGTTT CTTGAAGATG AAAGCATATG
17001 GTAAGATGAA ATAGTGATTA TTTTTTAAAA ATTACTACTC CAGAAAGGAA
17051 AAGTTTACTA ATTTTATTA CTAAAGTTTA CTGTTGGTGG GTGCGGTGGC
17101 TCACACCTGT AATCCCAGCA CTTTGGGAGG CCGAGGCAGG CGGATCACCT
17151 GAGGTACAGG GTTCGAGACC AGCCTGACCA ATATGGTGAA ACCCCATCTC
17201 TGCTAAAAAT AAAAAATTAG GCCGGGCGCG GTGGCTCATG CCTGTAATCC
17251 CAGCACTTTG GGAGGCCGAG GCAGGTGGAT CAGGAGTCA GGAGATTGAG
17301 ACCACTCTGG CTAACACGGT GAAACCCGT CTCTGCTAAA AATACAAAAA
17351 TCAGCCGGGC GTCTTGGCAG GCACCTGTAG TCCAGGTAC TCAGGAGTTT
17401 TGAGACGGGA GAATGGCGTG AACCCGAAG GCGGAGCTTG AAGTGAGCCG
17451 AGATTGCGCC GCTTCAGTCC AGCCTGGACG ACAGAGTGAG ACTCTGTCTC
17501 TAAAAAATA AATAAATAA AATAAAAAAT TAGCTGGGTG TGGTGGCAGC
17551 CACCTGTACT CCCAGCTACT CGGGAGGCTG AGGCAGGAGA ATTGCTTGAA
17601 CCCGGGAGAT GGAGGTTGCG GGGAGCCAAG ATTGCGCCAC TGCACTCCAG
17651 CCTGGCGACA GAGTGAGACT CTTTCTCAGA AAAAAATATG ATAATTAAAA
17701 GTTGACACGT TCTTCGCCGA GAGTGGTCCG GGTTTCCTGC TTCAACAGTG
17751 CTTGGACGGA ACCCGGCGCT CGTCCTGCAC CCCGGCGGCG CGCCCATAGC
17801 CAGCCCTCCG TCACCTCTTC ACCATGCCCT CGGACTGCCC CAAGGCCCCC
17851 GCCGCAGCTC CAGCGCCGCG TAGCCACCAC TGCCGCTGCC GCCGCCTCTC
17901 CTTAGTCGCC GGCATGACGA CCGCGTCTAC CTCGCAGGTG CGCCAGAACT
17951 ACCACCAAGA CTCAGAGGCC GCCATCAACC GCCAGATCAA CCTGGAGCTC
18001 TACGCTCCTT CCATTTACCT GTGCGTGGCT TACTACTTTG ACAGCGATGA
18051 TGTGGCTTTG AAGAACTTTG CCAAATACTT TCTTCACCAA TCTCATGAGG
18101 AGAGGGAACA TGCTGAGAAA TTGATGAAGC TGTAAGACCA ACGAGGTGGC
18151 CGAATCTTCC TTCAGGATAT CAAGAAACCA GACTGTGCGG GGAGAATGCG
18201 ATGGGAGAGC GGGCTGAATG CGATGGATTA CATTTGGAAG AAATTGTGCA
18251 TTTTGCATTA CATTTGGAAG AAAATGTGAA TCAGTCACTA CTGGAAGTGC
18301 ACAAATGGC CACTGACAAA AATGACCCCC ATTTGTGTGA CTTCATTGAG
18351 ACACATTACC TGAATGAACA AGTGAAGGCC ATCAAAGAAT TGGGTGACCA
18401 CGTGACCAAC ATGCACGAGA TGGGAGCGCC CGAATCTGGC GTGGCAGAAT
18451 ACCTCTTTGA CAAGCACACC CTGGGAGACA GTGATAATGA AAGCTAAGCC
18501 TCAGGCTAAT TTCCCATAG CCATAGGGTG ACTTACCTTG TCACCAAGGC
18551 AGCGCATGTA TGTGGGGTT TCCTTTACCT TTTCTATAAG TTGTTCCAAG
18601 ACACCCACTT AAGTTCTTTG ATTTGTACCA TTCCTTCAAA TAAATAAATT
18651 TGGTACCCTC CCCCCCCCCA AAAAAAAAT GTACTGTGGG CTGGCGTAGT
18701 GGCTCATGCC TAAATCTCAG CACTTTGGGA GGCTGAGGCG GGAGGATCAC
18751 CTGAGGTGCG GAGTTTGAGA CCAGCCTGGG CAACATGGTG AAACCCGTG
18801 TCTACTAAAA ATATAAAAC TAGCCAGTCA TGGTGGCACA CACCTGTAAT
18851 CCCGGCTACT TGGGAGGCTG AGGCATGAGA ATCACTTGAA CCTGGGCTGC

```

FIGURE 3, page 6 of 17

```

18901 GGAGGTTGTA GTAAGCTGAG ATCATGCCAC TGTACTCCAG CCTGGGTGAC
18951 AGGGAGACAT TCTCTCTCTC AAAAAAAAAA AAAAAACAAA AAAAAACAAA
19001 CAAACCAACA AAACAAAGTA ATCCAGGAAC AACAACATGA TGAAGGACTG
19051 CATGCAGGAC TCAGTGATGG ATGGTGGAAG ACAGCCAGGA AGTTAAGCAT
19101 GACTCTGGTA TTAAGTGTG TCTGGGAGAG TTAAGATTCC ATTTACAGAA
19151 ATAAGACCTG TAGGGGAAGC TCTTGATTTT TTTTTTTTGG CAGACTGCTG
19201 ATTTCTGAT TACATGTGTT AAGTTTGAGG TATAGAGAGA AAGAACATCC
19251 TGGCCGGGTG CAGTGGCTCA CACCCGTAAT CCCAGCACTT TGGGAGGCCA
19301 AGGTGGGCAG ATCACGAGGT CCAGGAGATC GAGACCATCC TGGCCAACAT
19351 GGTGAAACCC CGTCTCTACT AAAAAATACAA AAATTAGCTG GGCGTGGCGG
19401 CGCGTGCCCTG TTATCCCAGC TACTCAGGAG GCTGAGGCAG GAGAATTGCT
19451 TGAACCCGAG AGGCAGAGGT TGTGATGAGC CGAGATCGCG CCACTGCACT
19501 CCACCCTGGC AACAGAGCTA GACTCTGTCT CAAAAAATAA AAAAAGAAAG
19551 AAAAAAAGAA CATCCTGTAG AAACAGGCAG TCAGAGGTAT AGAACTACAC
19601 AGAATCCAAG AGATCTTTCA AGAAAAGTGA CATGCAGCAA GAGAAACTAT
19651 CAAGGGGGTA AACAACCTAT AGAATGGGAG AAAATATTCA CAAAGTATAC
19701 ATCCAAACAA AGTCTAATAT CCAAAATCTA TAAGGAACCT AACAAAGCAA
19751 AAGCAAATAA CCCCCTTAAA AAGTGGGCAA AGGACATGAA CAGATACTTC
19801 TCAAAAGACG TACATGTGGC CCACAAACAT GAAAAAACGC CCATTTCTAA
19851 TCATCAGACA AATGCAAATT AGAACCACAA AGAGATACCA TCTCACACCA
19901 GTCAGAACAG CTTTTGTAA AATGTCAAAA AATGAGAAAC GTTGGTGAGG
19951 CTGCAGAGGA AAGCAGACAC TTGTACACTG TTGGTGAAGG TGTAATTAAT
20001 TTTAGCGTAG GCACAGTCAG TTTGGAGATT TCTCAGAGAA CTAAGAGTGG
20051 AACTACCATT AGACCCAGCA ATCTGATGGC TGGATATACG GCCAAAGGAA
20101 AATAAATCAT TCTGCCAAAA GAACATATGT ACCTGTATGT TCATTGTGGC
20151 ACCATTACAC ATAGCAAAAA CATTGAATCA ACTCATGTGC CCATCAGTGG
20201 CGGACTAGAA AAGAAAAGAA AATATGGTAC ATAGCCATCA TGGAACTACT
20251 TGCACCCATT AAAAATAATG AAATAATGTC TTTGCAACAA CATGAATGTA
20301 GCTGGAGGGC ATTATCCTAA GCAAATAAC ACAGAAACAG AAAACCAAAT
20351 ACTGCGTGT CTCACGCAGT GAGAGTGGGA GCTGAACATC AAGTACACAT
20401 GGATGTAAGG ATGGCAACAA TAGACATGGG TCTACTAGAG GTGGTGGTGC
20451 GGCAGGGTGG GGGTGGGGGT TGTGTGGCAG AGGAACAGCT GAAAAACTAC
20501 CTATTTGATA CTATACCCAG CACCTGGGAA ACGGGTTCAG TCATACCCCA
20551 AACCTCAGCA TCACACAGTA TACCTTTCTA ACAAACCTAC ACATGTATTC
20601 TGTGATTCTA AAATAAACAT TGAAAAATAA AAAAAAACT GACATGGTTT
20651 GTACTGTTTA ATCTGACATA ATGGCTAGGG GAAATGAAGT CTGCAGAATG
20701 GCTGTTTACG GATGTTGTG TTGTTGTTGA GATGAGGTCT CACTATGTTG
20751 CCCAGGCTTG AACTCCTGGC CTCAAGCAGT CCTCCTGCCT TGACCTCCCA
20801 AAATGTTGAG ATTACATGCA TGAGCCATTG CCAAACGGC TATTTGGATT
20851 GCTGTTAAGG TTATTACATT CTCTGTGTAG TAAGACCTTG AAGGAGAAGG
20901 ATTTGAGATC AGGAGTTTAA GAAAAAATGT TAATCTAGGA AGAGAGGATA
20951 ATTTCTGTGT TTGGCCAGTT GCAATGGCTC ACGCCTGTAA TCCCAGCGTT
21001 TTGGGAGGCC GAGCTGGGCA GATCACTTGA GCTCAGGAGT TTGAGAAGAG
21051 CCTGGGCAAC ATGGTGAAGC CCCGTCTCTA CTGAAAATAC AAAAATTGGC
21101 TGGGCGTGGT GGCAGGTGCC CACTGTAATT TCAGCCACTC AGGAGACTGA
21151 GGCAGGAGAA TTGCTTGAAC CTGGAAGTG GAGGATATGG TGAGCCTAGA
21201 TCGCGCCACT CTACTCCAGC CTGGGAGACT CCATCTCAA AAAAAAAGAA
21251 AACAAAGATG TGAAATGAAG TAATTACCAC AGTCAATGTG ATCCTATAAC
21301 TTTGTTTTCT TTTAGAGATG GGGTCTCCCT CTGTCACCCA GGCTGGAGTG
21351 CAGTGGTGCA TCATAGCTTT CTGCAGCCTC CACCTCCTGG GCTCAGGTGG
21401 TCCTCTTGCC TCAGTGTCC GAGTAGTTAG GACTGACTGC AGGTGCATGC
21451 TGCTATGCCT GGCTAACTTT AAAATTTTTT TGTAGAGGCG GGGTCTTGCT
21501 ATGTTGCCTA GGCTGGTCTC CAACTCCTGA TCTCAATCAG TCCTCCTGCC
21551 TACCTTCCCA GAGCGCTGGG ATTACAGGTG TGAGCCATCG CACCTAGCCA
21601 ATCTCATAAC ATTTTATGAC TAGCAAACCT AGTAGTTCTG ATTCAGGCAT
21651 AAATCAGTTG GTGGGGTTAT ACAAGGTTGG GTGAGTTTTT CTAGATTTCT
21701 AAGAGACCAT GTTGAAATAC TTGGCTCTGT CTCAGTAAGG GATAGAGAGA
21751 AGCAAAGGTG TGGGTAAAGG TTATGAGCAG ACATGTAAAG GGGGCAAAAT
21801 AAAGTGTTTA GGAAGGTGA AACAGTTCCA AATCATAATA TAGATCCCAA
21851 GCCTACCAA GAAGTGAAGG AGAGGGTAAT TGTGCGATAG TTCTCAGGAC
21901 TGAGACCTCA AGGTGTAAGA CGAATCTTTA TTGTGGGTGG TCCTCTTTGA
21951 GAAAAAAGAA AAAAAAGAAA AATGTGAAAT GGGTGCTAAA TTTCAGAGCA
22001 GAATATACAT ATATGTGTAT GTGTGTGTGT ATATATATAT ATATTTTTAT

```

FIGURE 3, page 7 of 17

22051	GTAGCAAAATT	CCAGTTTACA	AAGGGCTTTT	ATGTGTTTGT	CATCATTACA
22101	ACAGTTCTGT	GATGATGTGG	ATGTGGTGGA	TATTGTAATT	CCACATCCCA
22151	GAGGATAAAA	TTGAGGCACA	CCATAAGGTA	GCTGACCAGA	GATCATGCAG
22201	TATATGGTAG	AATGGCAACT	TGAGGCCAGT	TCTTAGGTGT	GTTTGTGTGC
22251	CTTATTCAGA	ATGGAACAAT	GTGGTTTATT	GTACAAAAAT	TTAAAAATGA
22301	ATGTCGAAAA	GTAGAAATTA	TTACCCAAAT	CTCACCTCT	GTGGTTGTTT
22351	GCTATGTGAT	CTTCCAGACT	CACATATACA	TGTAGATAAT	TTTTTTTTTT
22401	TTTTTTGAGA	CGGAGTCTTG	CTCTGTCGCC	CTGGCNNNNN	NNNNNNNNNN
22451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
22951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
25001	NNNCAAAAAT	TAGCTGGGCG	TAGTGGTGGG	CGCCTGTGGT	CCCAGCTACT
25051	TGGGCAGCTG	AGGCAGGAGA	ATCGCTTGAA	CCTGGGAGGT	GGAGGTGCA
25101	GTGAGCCAAG	ACCATGCCAC	TGCACTCCAG	CCTGGGCAAC	AGAGCAAGAC
25151	TCTTGTCTTA	AAAAACAAAA	AAGTGATACC	AGTTGAGCTG	ATTCTTTATC

FIGURE 3, page 8 of 17

```

25201 TTTTTCAC TGGAGAACTA AGTATACAGG TGAGAAAAGA CGAGATATTT
25251 ATACCCGAGA GAATTGATGG TGAAATCCAT TTTTGGGAT CAGAACTTCC
25301 CCAAACAGTG TCCTTCAAAT AGGGTTCAGG GGTGCTAAGA TATTTATCCC
25351 CTCAACCCTT GGGGTTCAC TCCAGTATGGC ATATAAATAT TGTATCACTT
25401 TCTATGTGTG GGGAGCAGTG CTCCAGGTGA CCTTCCTTCC TTTCTTCTA
25451 GGGGAGGAAG TTGCCTTGTA CTGTGCCAAA TATCTTCTCG ATATCATCAA
25501 AGATCAGAAG GCCTACAAGG AAGGCAAGCT ACAGAAGGTC TGTCTGCTTA
25551 CACCGCCCAT TCCTCACTTG TGTAGGCTTT TCCCTTGTTT TCTAGCCCTT
25601 GGGCTTTTCC TTTCTTTTGG TCCTCTAGCT GCTGCTGCTT ATTTACTCTT
25651 GAAGAATTCT GTTCCTAAAA CGAGCTTATT GGCCGCCTTT TAGACTTGCC
25701 TTATTATTCC TAGGCCTCTG AGCTGTTTTT ATCTGTGAGT GTCTCTTAGT
25751 GTGGTGGCTC AACTCTTAA TTTGTATTCC ATCCTTGTCG TCAGGATTGT
25801 ATATAGGGAG TTCATTTTGT ACTAGTCTTA GACTATTTTG CTTATATTCA
25851 GGCTTTAGAA GATGCCTTCT TGGCTATTGA CGCCAAATTG ACCACTGAAG
25901 AAGTCATTAA AGAGCTGGCA CAGATTGCAG GGCGACCCAC TGAGGATGAA
25951 GATGAAAAAG AAAAAGTAGC TGATGAAGAT GATGGTGAGT GTGGCATCCC
26001 TTGTTTGAGG GGAATCAGC ATTTTAAGAA ATATTCTTTA ATATTACTTA
26051 TCAATTCTAA GATAGGATGG CTTTCTAGGG ACCTGGGGAG TCCTTATGTT
26101 AAAGAAACCT ATGATGTTCT CCTGCATTGT ATGTGGTTAT GAAAAGGAGG
26151 GAGAGAATTA TCTTCGTGA GTGGCATCTG AGCTGTAAGC ATTGTATATA
26201 CATTATCTTT TGTCACTGTG ATGGGGTCTT CCTGGTTCCT GCTAGTATTT
26251 ATGTGCTTTT TTTTCCCTC AAGACTGGAG CAGTTATTAG CCCCATAGC
26301 CAATCATTA GCCTAAATCC TAATCACAG TAGCATTGTG GGCTTCCTGG
26351 ATCCTCAGCC AGAATAGGGT TTTTACAAC TAACAATAAA AAATGAGACG
26401 TCAGAGGGGA AGTATAGTAA CTAGTGTGT TTTGATTAG AAGGGGATGA
26451 AACACAAAAA CCAAAGAAG TCTGTGGAGG AGGAGGAGCT AGGGCATGTT
26501 CTTCTGAGAC TTGAGCGAGA GGAACCTTGG GAGTGGGAGG TTGTGGGGAA
26551 GTTAGAGGCT GCAAGGGCTG TTGAGGTAGT GAGAGGGACG GATCCCATGA
26601 GGAGTCTGGC ATGGGGGCTC TGATTTAGCC TCTTCCCTGC AGTGGACAAT
26651 GAGGAGGCTG CACTGCTGCA TGAAGAGGCT ACCATGACTA TTGAAGAGCT
26701 GCTGACACGC TACGGGCAGA ACTGTCACAA GGGCCCTCCC CACAGCAAAT
26751 CTGGAGGTGG GACAGGCGAG GAACAGGGT CCCAGGGCCT CAATGGGGAG
26801 GCAGGACCTG AGGACTCAAC TAGGGAAACT CCTTCACAAG AAAATGGCCC
26851 CACAGCCAAG GCCTACACAG GCTTTTCTC CAACTCGGAA CGTGGGACTG
26901 AGGAGGCCA AGTTGGTGAG CCTGGCATT CCCTGGTGA GGCTGGGCTT
26951 TCCTGCTCTT CAGCCTCTGA CAAGCTGCCT CGAGTTGCTA AGTCCAAGTT
27001 CTTTGAGGAC AGTGAGGATG AGTCAGATGA GGCGGAGGAA GAAGAGGAAG
27051 ACAGTGAGGT AAGGGCCTGT GAGGGCAGGC AGATGCTGAA GTTGACAGGA
27101 GGTCTGTTT GGTGCGCTG TGTAGTTTT AACTCTCTT CCTTCTCTA
27151 TTTTGACATT ATCCCCAAG ACCCACTGTA TTCTAAGCTT TAGTCTTGAA
27201 TTCATTGAGC TCCATCATCA CAGGTACCAT TTGCCTTTT ACCTCTTCTT
27251 TTGTTGGTAC TATAACAAGC AGATCTAGT CTGGCTTTT AGAGTCTGTC
27301 TCCTAGAGAG AGAACAAGGA GATAGTTGTT ACCTTGGCTA GTTGACTGTT
27351 TTCTTCTCTG GAAAATTAT TTTCTGGCCA CAGTGCCTGA AAGATATTTT
27401 TGGCTGGCAG CCCTTGCTT GTCTGGGCT TTTGCTAGT GACTGCTAAG
27451 CCCAGTTTCA GATGTCAGT GTACTCATG TAGCCCTTC CATCCCCCA
27501 ATTTTCATGA CCATATACTT GTATCTTTCA GTGTTTTGAG GACCTGTGTT
27551 CAGTCAGGAC CTCTTGATTC TGAGTATGAG CTGTGGGGAG GGAGGGGATC
27601 ATCCCAGTCT CAGCAGTCTG GGATCCTCC CCTGGCAGGA ATGCAGCGAG
27651 GAAGAGGATG GCTACAGCAG TGAGGAGGCA GAGAATGAGG AAGATGAGGA
27701 TGACACCGAG GAGGCTGAAG AGGACGATGA AGAAGAAGAA GAAGAGATGA
27751 TGGTGCCAGG GATGGAAGGC AAAGAGGAGG TGTGTGGGGA AGGGGAGCAA
27801 TGAGTCTTGA AAAGCCACAA GGCAGGTGTG AATCCCTTAA TTTTGATTTT
27851 GAGACAGGGG ATCCCCCTGA TACTTTAGGA TGGAAGTAAT AGTCATGGGG
27901 ATTTATTCTG CAAGGGGAAT GAGATGGTAA GCCTTTGGGG TTGAATTATC
27951 TAAAAACAAG GGAGAGGGAG TGTGCTGCTG TCTCTAGAAA GATGAAATGT
28001 GTGCTTCTCC TGTGTTGTTA AGCTCTTTT GGGGTCCCAG TGAAAGCAAG
28051 CATAGGTGAA CGATCAGGAG CACATCAGTG AGGAACGCAT GTTCAGAAGC
28101 CCCCATGATG CTCCTTTTCT TCCTCTTAAG CCTGGCTCTG ACAGTGGTAC
28151 AACAGCGGTG GTGGCCCTGA TACGAGGGAA GCAGTTGATT GTAGCCAACG
28201 CAGGAGACTG TCGCTGTGTG GTATCTGAGG CTGGCAAAGC TTTAGACATG
28251 TCCTATGATC ACAAACCAGA GGATGAAGTA GAACTAGCAC GCATCAAGAA
28301 TGCTGGTGGC AAGGTCACCA TGGATGGGCG AGTCAACGGG GGCCTCAACC

```

FIGURE 3, page 9 of 17

```

28351 TCTCCAGAGC CATTGGTAAG GGCCAAGAAA CTGGGAAAGA NNNNNNNNNN
28401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNATGGTGAA ACCCCATCTC
28751 TACTAAAAAA AAAAAAAAT AAAAAATTA GCTGGGCATG GTAGTGCAAG
28801 CCTGTAATCC CAGCTACTCA GGAGGCTGAG GCAGGAGAAC CGCTTGAATC
28851 CGGGAGGCGG AGGTTGTAGT GAGCCGAGAT CGTGCCATTA CACTCCAGCC
28901 TGGGCTACAA GAGTGAAACT CCGTCTCAA AAAAAAAAA CAAAAAGAC
28951 TTAAATAAAA AAGACCAGTG AGTGACTTTC TTAAGGTTCA GCAGTCTGGT
29001 GGCAGGGTTG AAAC TAGAAA AACTAGGACT TAGGACTCAG TTCCCCATTC
29051 CACTAGATTA TGGAACTTTG TAAAGAAGGG AAATGAATGG CAAGGTTTGA
29101 CCTGCCACAA ACACAAGTCT GTGGGAAGTA TCCAACTGC TCATCAACCA
29151 TTCCTTTACT CCAGGGGACC ACTTCTATAA GAGAAACAAG AACCTGCCAC
29201 CTGAGGAACA GATGATTTC GGCCTTCCTG ACATCAAGGT GCTGACTCTC
29251 ACTGACGACC ATGAATTCAT GGTCAATGCC TGTGATGGCA TCTGGTGAGC
29301 ACTGGCAGAA TGCCCTAAAT TCCCCTTTCT GCAGCATGTC TTCTCTATA
29351 GGACTCAGGG CACCTCTAGG ATTAGAGCCT AGGCAGACCT AGGCCTCTTG
29401 GTGGGTGAAG AGCACCCAGA CTAAGGCAGA GCTGAGAATT TCTGTAGTTA
29451 TTTACACTGG CCTGGGCCAC CACCTCTGTC CATACTCCTC TACGCTGCCT
29501 TAGTGAGACT GGAAGATTCT GACTGTTGTT CTTGACCCCA GGAATGTGAT
29551 GAGCAGCCAG GAAGTTGTAG ATTTCAATCA ATCAAAGATC AGCCAGCCTG
29601 ATGAAAATGG GGAGCTTCGG TTATTGTCAT CCATTGTGGA AGAGGTGAGT
29651 ACCAGGGTGG AGAAGAGAGG GTGTCTGGTC TGCACAGCCA GGGTT (SEQ ID NO:3)

```

FEATURES:

```

Start:      2134
Exon:       2134-2326
Intron:     2327-25451
Exon:       25452-25537
Intron:     25538-25851
Exon:       25852-25984
Intron:     25985-29164
Exon:       29165-29294
Intron:     29295-29541
Exon:       29542-29647
Stop:

```

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA

Position	Major	Minor
2944	A	-
3232	G	A
7236	T	A
8187	T	A G
8187	T	- G
8187	-	A G
9172	C	A
9173	T	-
9180	T	-
13351	T	C
15100	G	C
15255	A	C
18996	A	C
18996	A	C
18996	A	C

20532	C	T
21118	G	T
21211	C	G
21212	T	C
21238	-	A
26483	G	A

Context:

DNA

Position

2944

AGATGTGAATAACTTTTTAAATTTTAATATTTAAAATATCTGATGTGGGAAGCCTCTTTT
GGCTAGGAGTTTGACAGTGAAAGGAACCCCGGCAGAGTCTGTTTCACATTTTGGTTGCC
TGGCCTTGGGCTCTTGCAATTTTCAAGAGGCTGGACCCGACCTCCAGGAGTTGTAC
TCATTTGCACTCTTTTCAAGGCTTTTACTAACTTCGGAAGCTTGAATTATGTCAGTCC
CTAGGTTTTCCTTTTATTATGCTTCGTTTCTTTCTCTCTTAGGATTTCTCTAAACTT
[A, -]
ATCAGTAATTTCTCATTTGCCTGTAATTGTAGTTTATAGTTTCTTTGATTGTAAGTCATTT
CATCAATTTTTTTCATCACAACTACCTACAAAGGGCTTTCTAGAAAATTTTACTCTGG
ACAAAGGGGAAAAGAAAAATATTGGGGGAAAGTAGTAGTATTAGGTAAACTTGATGT
GAAACTACAAAGAGAAGAGGGAAAAGTGCAGTAGGGAGGAAAGGAGGAAGACGGGTTA
ACCGTGGCTTTGTGAAGAGCATTCTGAAGTCTAGGCAAAAGGGCCAGGAAAATCTCTGT

3232

TCTCTAAACTTAATCAGTAATTTCTCATTTGCCTGTAATTGTAGTTTATAGTTTCTTTGA
TTGTAAGTCATTTTCATCAATTTTTTTCATCACAACTACCTACAAAGGGCTTTTCTAGAA
AATTTTACTCTGGACAAAAGGGGAAAAGAAAATATTGGGGGAAAGTAGTAGTATTAGG
TAAACTTGATGTGAACTACAAAAGAGAAGAGGGAAAAGTGCAGTAGGGAGGAAAGGGA
GGAAGACGGGTTAACCGTGGCTTTGTGAAGAGCATTCTGAAGTCTAGGCAAAAGGGCCAG
[G, A]
GAAATACTCTGTCTGGTATTGAGGGTTTCTCCACCTACCGGGTGGGCTTCAGGTAACAGC
GAAATACTGTCTCCCTTGGGAATTGTTTCAGATCCCTCGCTCCTGTGGTTAGCTCTG
GAATGCCAGTATGAACCTCAATGTTTGTGTTTTCGATTCAAATTTTATATTCTAAGTGA
CCTTAATAACAATTTTACAATTAGGTATAAAATTTTCAGGATCCTAGTGTATCCTATAGTT
CATCTCATCTGCTTTGGCTCCCTTTTTTTTTTTTTGAGACAGAATTTGCTCTTGTGGC

7236

GACAGGATTTCACTATGTTGGCCAGGCTGGTCTCAAGACTCCTGACGTTGTGATCCACCT
GCCTCGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCTACCGTGCCCGGCCAGGAATT
TTTTGTGCTATAAATCATATTTTCTTTTATTAAAGGCAGTGTCAATATCTATAGTATAA
TTTTGAGGAGGCTGGCTATTATTGCTGTGTAGAAGCTGGCTTATTAGTGGTCAAGGGGT
CATCTAGAATTGACTATAAAGATAGTATTGAGCAGAAAATTTCTTAAATACCTGCATATT
[T, A]
GTTTCAGTCATTAAATTAATGGAAAAATATAAAAAGAAATATCACAAAGTATGCTATGGG
GTTCTACCTTAGGGCATTGAAGGTTGAAAAACATTTATTTTCTTATCTTCAGATTAGCAT
CTCATATCAATAGCCAATAGCTTAAAGCGCTTTTACTTACTAAACCAGGTCAGAATTTCT
TCTCTCTCTCTTTTTTTTTTTTTTAGACAGAGTCTCTGTCACCCAGTCTAGAGTGCAGTG
GTGTGATCTCGGCTCACTGGCCTTTCTGTTTAAAGTATTCTCGGCCTCAGCCTCCTGAG

8187

GTTGCCAGGCTGGAGTGCAATGGTGCCATCTCGGCTCACTGCAGCTTCTGCCTCCTGGG
TTCTAGCAATTTGCCTACCCAGCCTCCCAAGTAGCTGGGATTACAGGCACTCGCCACCA
TGCCAGCTAACTTTTTTGTATTTTGTAGTAGAGAAGGGGTTTTCCCATGCTGGCGAGTCT
GGTCTTGAACCTCTGGCCTCAAGTGATCTGCCCGCTCAGCCTCCCAAAGTGCTGGAATT
ACAGGCATGAGCCACTGTGCCCGGCCTTTTTATTTTTTAAATTTATTTATGATTTATTTT
[T, A, G]
AGACAGGATCTCACTCTTGCCCATGCTTGAGTGGTAAGAGTATGGGATTTCTTGTGCCG
GTTCCCTCACATCCTCACTACACTTATCTGCCTTACAGTGGCTCGATCATAGTTCACTG
CATAGCCTTCTGGGCTCAAGGGGTCTCCAGCCTCAGCCTAATATAGGCACATGCCACCA
TGACTGGCTAATTTTTTTTTTTTAAAGTTTTTTTTTGTAGAGATAGGGCCTTGCAGTGTG
CCCAGGCTGAGGAATTTTATTTATGTTTATTTTATTTATTTATTTATTTATTTATTTATTT

8187

GTTGCCAGGCTGGAGTGCAATGGTGCCATCTCGGCTCACTGCAGCTTCTGCCTCCTGGG
TTCTAGCAATTTGCCTACCCAGCCTCCCAAGTAGCTGGGATTACAGGCACTCGCCACCA
TGCCAGCTAACTTTTTTGTATTTTGTAGTAGAGAAGGGGTTTTCCCATGCTGGCGAGTCT
GGTCTTGAACCTCTGGCCTCAAGTGATCTGCCCGCTCAGCCTCCCAAAGTGCTGGAATT

FIGURE 3, page 11 of 17

ACAGGCATGAGCCACTGTGCCCGGCCTTTTTATTTTTTAAATTATTTATGTATTTATTTT
[T, -, G]
AGACAGGATCTCACTCTTGCCCATGCTTGAGTGGTAAGGAGTATGGGATTTCTTGTCGCCG
GTTCTTCCACATCCTCACTACACTTATCTGCCTTTACAGTGGCTCGATCATAGTTCACTG
CATAGCCTTCTGGGCTCAAGGGGTCTTCCAGCCTCAGCCTAATATAGGCACATGCCACCA
TGACTGGCTAATTTTTTTTTTTAAGTTTTTTTTTGTAGAGATAGGGCCTTGCAGTGTG
CCCAGGCTGAGGAATTTATTTATGTTATTTTATTTATTTATTTATTTATTTATTTATTT
8187 GTTGCCAGGCTGGAGTGCAATGGTGCCATCTCGGCTCACTGCAGCTTCTGCCTCCTGGG
TTCTAGCAATTTGCCTACCCAGCCTCCCAAGTAGCTGGGATTACAGGCACTCGCCACCA
TGCCAGCTAACTTTTTTGTATTTTGTAGAGAAGGGGTTTTCCCATGCTGGCGAGTCT
GGTCTTGAACCTCTGGCCTCAAGTGATCTGCCCGCCTCAGCCTCCCAAGTGCTGGAATT
ACAGGCATGAGCCACTGTGCCCGGCCTTTTTATTTTTTAAATTATTTATGTATTTATTTT
[-, A, G]
AGACAGGATCTCACTCTTGCCCATGCTTGAGTGGTAAGGAGTATGGGATTTCTTGTCGCCG
GTTCTTCCACATCCTCACTACACTTATCTGCCTTTACAGTGGCTCGATCATAGTTCACTG
CATAGCCTTCTGGGCTCAAGGGGTCTTCCAGCCTCAGCCTAATATAGGCACATGCCACCA
TGACTGGCTAATTTTTTTTTTTAAGTTTTTTTTTGTAGAGATAGGGCCTTGCAGTGTG
CCCAGGCTGAGGAATTTATTTATGTTATTTTATTTATTTATTTATTTATTTATTTATTT
9172 AAGGATATTTACACAAAGTTCTATTGTCTGATTCCCTAGCAGTTGTTACTACTGTTTCC
CTAACCTCTAATCTTCTATTGGGTTATTAGTCTTAGAATTGAATTTGAGAGGTAAGGGC
TTGAATTTGAACATAGAAATTTATACAGGCTGATCAGTAGTTCTTGACATTGTATTATC
TGGAACAAATCTTTAGAACTGAGCTTAAGATGTTAATGACATTTTGTAGACAGAGTAT
GATTTCACTGATGTTGTTTTGTTTCTTTCTAGATCTAGTTCAGAGATGAAGTATATCAA
[C, A]
TTTTTTTTCTTTTTGACCCAATGCTAGCAGAAAAACAACACCTTTTAATCATATTTAGT
ATTTGAAAATGTGTATACAGGTTCCCTTTTTATTTATTTATTTCTTACAGGTTCCCTTTTT
AATCAGCTTTATTGAGATAGAGTTCATATACTGTATGGTTCATACCACATATGGTTCATA
TACCATACAGTATATGAACCTCACTTTAAGAGTATAATTCACTGGGTTTAAGGGTATAATT
CATTCAATTTAAGGGTATAATTCACTGGGCTTTTAGTATATTTCTTTTTTTCTTCTTTT
9173 AGGATATTTTACACAAAGTTCTATTGTCTGATTCCCTAGCAGTTGTTACTACTGTTTCCC
TAACCTCTAATCTTCTATTGGGTTATTAGTCTTAGAATTGAATTTGAGAGGTAAGGGCT
TGAATTTGAACATAGAAATTTATACAGGCTGATCAGTAGTTCTTGACATTGTATTATCT
GGAAACAAATCTTTAGAACTGAGCTTAAGATGTTAATGACATTTTGTAGACAGAGTATG
ATTTCACTGATGTTGTTTTGTTTCTTTCTAGATCTAGTTCAGAGATGAAGTATATCAAC
[T, -]
TTTTTTTCTTTTTGACCCAATGCTAGCAGAAAAACAACACCTTTTAATCATATTTAGTA
TTTGAAAATGTGTATACAGGTTCCCTTTTTATTTATTTATTTCTTACAGGTTCCCTTTTTA
ATCAGCTTTATTGAGATAGAGTTCATATACTGTATGGTTCATACCACATATGGTTCATAT
ACCATACAGTATATGAACCTCACTTTAAGAGTATAATTCACTGGGTTTAAGGGTATAATT
ATTCATTTTAAGGGTATAATTCACTGGGCTTTTAGTATATTTCTTTTTTTCTTCTTTT
9180 TTTACACAAAGTTCTATTGTCTGATTCCCTAGCAGTTGTTACTACTGTTTCCCTAACCTC
TAATCTTCTATTGGGTTATTAGTCTTAGAATTGAATTTGAGAGGTAAGGGCTTGAATTT
GAACATAGAAATTTATACAGGCTGATCAGTAGTTCTTGACATTGTATTATCTGGAAACA
AATCTTTAGAACTGAGCTTAAGATGTTAATGACATTTTGTAGACAGAGTATGATTTCACT
TGATAGTTGTTTTGTTTCTTTCTAGATCTAGTTCAGAGATGAAGTATATCAACTTTTTTT
[T, -]
CCTTTTTGACCCAATGCTAGCAGAAAAACAACACCTTTTAATCATATTTAGTATTTGAAA
ATGTGTATACAGGTTCCCTTTTTATTTTATTTATTTCTTACAGGTTCCCTTTTTAATCAGCT
TTATTGAGATAGAGTTCATATACTGTATGGTTCATACCACATATGGTTCATATACCATAC
AGTATATGAACCTCACTTTAAGAGTATAATTCACTGGGTTTAAGGGTATAATTCACTCAT
TTAAGGGTATAATTCACTGGGCTTTTAGTATATTTCTTTTTTTCTTCTTTTTCTTTT
13351 TTTTCCCCCATGAATTATCTTGGCATCATTGTCAGAGATAAATTGACCGTAAATGTGAG
GGTTTTATTTCTGAACCTCTCAAGTCCATTTTATTGGTCTACATGTCCCTATGCCAGTAAT
ACACTATCTTGGTTACTGTAGCTTTTTTAGTACGTTTGAATGTTTTTAAATTTGTTTT
TCATCTAAATTTTAGGATTAATTTGTCAATTTCTGCACAAAAGGCACCTGGGTTCTATA
GGGGTTATGCAGAATCTGTAGATCACTGGGGAGTATTACAGGCATGAGCCACCGTGCC
[T, C]
GGCTGACTGAGTTTTTCATAGATGTACTCTATCAGGTTTAGGAAGTTCCTTTTTATTCTT

FIGURE 3, page 12 of 17

AGGTTGTTGAGTCTATTTTATATTACTTTTTTAGAGACAGTCTTGCTCTGTCCCTCAGGC
 TGTAGCACAGTGGCTCAATCATAGCTCACTGCAGCCTTGAACCTCTAGGTTCAAGAGATC
 CGCCTGCCCTCAGCCTTCTTAGTAGCTGGGATTACATGCATGCACCACCATACTGGGCTAA
 TTTTTTAAATTTTTTATAGAGACAGGCTTATTACTATGTTGCCAGACTGGCATTGA

15100 GTTGTGTACCCAGTTTCCAGAACTCTTCATCCTACAGAACTGAAACTCCATACCCATTAA
 ATGAGTCCCCATTCTCTTTCCCCAGCTCATGGCAAACAGCATTCTATTTTCAGTCTCTA
 TGAATTTGATTAGTTTATAGATACTTCATACTGTAAGTGGAAATCATATGGTATTGTCTTTT
 AGTGACTGCCATAATTTAAAAAAATTTTTTTGAGACGGAGTCTGCTCTGTGCCCCAGGC
 TGGAGTGCAGTGGCACCATCTCTGCTCACTGCAACCTCCACCTCCCAGGTTCAAGTGATT
 [G, C]
 TCCTGCCTCGGCCTCCACGTAGCTGGGATTACAGGTGCTCGCCACAACACCCGGCTAAT
 TTTTGTATTTTATAGGTAGAGACTGGGTTTACCATTGTTGGCCAGGCTGGTCTCGAACTCC
 TGACCTCAAATTATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC
 ACTGTGCCAGCCTCCATGTTGTTTTTACAAACACCTGTATCATTTACATTTCCACCAAC
 AGTACACAAGAATTTAGTTTCTCCACATCCTTGCTAGCAGTTGTTATTATCTGTTTTTT

15255 TGGAATCATATGGTATTTGTCTTTTAGTGACTGCCTAATTTAAAAAAATTTTTTTGAGA
 CGGAGTCTGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCACCATCTCTGCTCACTGCAAC
 CTCCACCTCCCAGGTTCAAGTGATTGTCTGCTCGGCCTCCCACGTAGCTGGGATTACA
 GGTGCTCGCCACAACACCCGGCTAATTTTTGTATTTTATAGGTAGAGACTGGGTTTACCA
 TGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAATTATCCACCTGCCTTGGCCTCCCA
 [A, C]
 AGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCAGCCTCCATGTTGTTTTTCAACAAC
 CTGTATCATTTACATTTCCACCAACAGTACACAAGAATTTAGTTTCTCCACATCCTTGC
 TAGCAGTTGTTATTATCTGTTTTTTTTTAATGGTTTCTTTTTTCTTTTTTCTTTTTTTT
 TTTTTTGAGACGGTCTTATTCTGCTCATGCTGGAGTGCAGTGGTGAATGTGATAGCTC
 ACTGCAGCCTCAACCTCTGGGCTCAAGCAGTCTGCCACCTCAGCCTCCACATAGGTGGG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA
 GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA
 GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

18996 GTAGTGGCTCATGCCTAAATCTCAGCACTTTGGGAGGCTGAGGCGGGAGGATCACCTGAG
 GTCGGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCGTGTCTACTAAAAATATA
 AAAACTAGCCAGTCATGGTGGCACACACCTGTAATCCCGGCTACTTGGGAGGCTGAGGCA
 TGAGAATCACTTGAACCTGGGCTGCGGAGGTTGTAGTAAGCTGAGATCATGCCACTGTAC
 TCCAGCCTGGGTGACAGGGAGACATTCTCTCTCAAAAAAAAAAAAAAAAAACAAAAAAA
 [A, C]
 AAAACAAACCAACAAACAAAGTAATCCAGGAACAACAACATGATGAAGGACTGCATGCA
 GGACTCAGTGATGGATGGTGGAAAGACAGCCAGGAAGTTAAGCATGACTCTGGTATTAAGT
 GTTGTCTGGGAGAGTTAAGATTCCATTTACAGAAATAAGACCTGTAGGGGAAGCTCTTGA
 TTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA
 TTTTTTTTTTTTGCAGACTGCTGATTTCTGATTACATGTGTTAAGTTTGAGGTATAGA

FIGURE 3, page 13 of 17

GAGAAAGAACATCCTGGCCGGGTGCAGTGGCTCACACCCGTAATCCCAGCACTTTGGGAG

20532 TAGCCATCATGGAATACTATGCACCCATTAAAAATAATGAAATAATGTCTTTGCAACAAC
ATGAATGTAGCTGGAGGGCATTATCCTAAGCAAATAACACAGAAACAGAAACCAAATA
CTGCGTGTCTCAGCAGTGAAGTGGGAGCTGAACATCAAGTACACATGGATGTAAAGA
TGGCAACAATAGACATGGGTCTACTAGAGGTGGTGGTGGCGCAGGGTGGGGGTGGGGT
GTGTGGCAGAGGAACAGCTGAAAACTACCTATTGATACTATACCCAGCACCTGGGAAA
[C, T]
GGGTTCACTCATACCCCAAACCTCAGCATCACACAGTATACCTTTCTAACAACTTACAC
ATGTATTCTGTGATTCTAAAATAAACATTGAAAAATAAAAAAACTGACATGGTTTGT
ACTGTTTAATCTGACATAATGGCTAGGGGAAATGAAGTCTGCAGAATGGCTGTTTACGGA
TGTTGTTGTTGTTGTTGAGATGAGGTCTCACTATGTTGCCAGGCTTGAACCTCTGGCCT
CAAGCAGTCCTCCTGCCTTGACCTCCCAAAATGTTGAGATTACATGCATGAGCCATTGCC

21118 GCATGAGCCATTGCCAAAACGGCTATTTGGATTGCTGTTAAGGTTATTACATTCTCTGTG
TAGTAAGACCTTGAAGGAGAAGGATTTGAGATCAGGAGTTTAAAGAAAAATGTTAATCTA
GGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTGCAATGGCTCACGCCTGTAATCCCAGC
GTTTTGGGAGGCCGAGCTGGGCAGATCACTTGAGCTCAGGAGTTTGAGAAGAGCCTGGGC
AACATGGTGAAGCCCCGTCTCTACTGAAAATACAAAAATTGGCTGGGCGTGGTGGCAGGT
[G, T]
CCCACGTAAATTCAGCCACTCAGGAGACTGAGGCAGGAGAATTGCTTGAACCTGGAAAG
TGGAGGATATGGTGAGCCTAGATCGCGCCACTCTACTCCAGCCTGGGAGACTCCATCTCA
AAAAAAAAGAAAAACAAGATGCTGAAATGAAGTAATTACCACAGTCAATGTGATCCTATA
ACTTTGTTTTCTTTAGAGATGGGGTCTCCCTCTGTCAACCAGGCTGGAGTGCAGTGGTG
CATCATAGCTTTCTGCAGCCTCCACCTCCTGGGCTCAGGTGGTCTCTTGCTCAGTGT

21211 AGGAGTTTAAGAAAAATGTTAATCTAGGAAGAGAGGATAATTTCTGTGTTTGGCCAGTT
GCAATGGCTCACGCCTGTAATCCCAGCGTTTTGGGAGGCCGAGCTGGGCAGATCACTTGA
GCTCAGGAGTTTGAGAAGAGCCTGGGCAACATGGTGAAGCCCCGTCTCTACTGAAAATAC
AAAAATTGGCTGGGCGTGGTGGCAGGTGCCCACTGTAATTTAGCCACTCAGGAGACTGA
GGCAGGAGAATTGCTTGAACCTGGAAAGTGGAGGATATGGTGAGCCTAGATCGCGCCACT
[C, G]
TACTCCAGCCTGGGAGACTCCATCTCAAAAAAAAAGAAAACAAGATGCTGAAATGAAGT
AATTACCACAGTCAATGTGATCCTATAACTTTGTTTTCTTTTAGAGATGGGGTCTCCCTC
TGTCACCCAGGCTGGAGTGCAGTGGTGCATCATAGCTTTCTGCAGCCTCCACCTCCTGGG
CTCAGGTGGTCTCTTGCTCAGTGTCCGAGTAGTTAGGACTGACTGCAGGTGCATGCT
GCTATGCCTGGCTAACTTTAAATTTTTTTGTAGAGGCGGGGTCTTGCTATGTTGCCTAG

21212 GGAGTTTAAGAAAAATGTTAATCTAGGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTG
CAATGGCTCACGCCTGTAATCCCAGCGTTTTGGGAGGCCGAGCTGGGCAGATCACTTGAG
CTCAGGAGTTTGAGAAGAGCCTGGGCAACATGGTGAAGCCCCGTCTCTACTGAAAATACA
AAAAATTGGCTGGGCGTGGTGGCAGGTGCCCACTGTAATTTAGCCACTCAGGAGACTGAG
GCAGGAGAATTGCTTGAACCTGGAAAGTGGAGGATATGGTGAGCCTAGATCGCGCCACTC
[T, C]
ACTCCAGCCTGGGAGACTCCATCTCAAAAAAAAAGAAAACAAGATGCTGAAATGAAGTA
ATTACCACAGTCAATGTGATCCTATAACTTTGTTTTCTTTTAGAGATGGGGTCTCCCTCT
GTCACCCAGGCTGGAGTGCAGTGGTGCATCATAGCTTTCTGCAGCCTCCACCTCCTGGGC
TCAGGTGGTCTCTTGCTCAGTGTCCGAGTAGTTAGGACTGACTGCAGGTGCATGCTG
CTATGCCTGGCTAACTTTAAATTTTTTTGTAGAGGCGGGGTCTTGCTATGTTGCCTAGG

21238 GGAAGAGAGGATAATTTCTGTGTTTGGCCAGTTGCAATGGCTCACGCCTGTAATCCCAGC
GTTTTGGGAGGCCGAGCTGGGCAGATCACTTGAGCTCAGGAGTTTGAGAAGAGCCTGGGC
AACATGGTGAAGCCCCGTCTCTACTGAAAATACAAAAATTGGCTGGGCGTGGTGGCAGGT
GCCCACTGTAATTTAGCCACTCAGGAGACTGAGGCAGGAGAATTGCTTGAACCTGGAAA
GTGGAGGATATGGTGAGCCTAGATCGCGCCACTCTACTCCAGCCTGGGAGACTCCATCTC
[-, A]
AAAAAAAAGAAAACAAGATGCTGAAATGAAGTAATTACCACAGTCAATGTGATCCTATA
ACTTTGTTTTCTTTAGAGATGGGGTCTCCCTCTGTCAACCAGGCTGGAGTGCAGTGGTG
CATCATAGCTTTCTGCAGCCTCCACCTCCTGGGCTCAGGTGGTCTCTTGCTCAGTGT
CCGAGTAGTTAGGACTGACTGCAGGTGCATGCTGCTATGCCTGGCTAACTTTAAATTTTT
TTTGTAGAGGCGGGGTCTTGCTATGTTGCCTAGGCTGGTCTCCAACCTCCTGATCTCAATC

26483 CTGTAAGCATTGTATATACATTATCTTTTGTCAATTGTGATGGGGTCTTCCTGGTTCCTGC

FIGURE 3, page 14 of 17

TAGTATTTATGTGCTTTTTTTTCCCTCAAGACTGGAGCAGTTATTAGCCCCAATAGCCA
 ATCATTAAAGCCTAAATCCTAATTCACAGTAGCATGTGGGGCTTCTGGATCCTCAGCCAG
 AATAGGGTTTTTACAACCTAACAAATAAAAAATGAGACGTGAGAGGGGAAGTATAGTAAC
 AGTGTGTTTTGATTAAGAAGGGGATGAAACACAAAAACCAAAGAAGTCTGTGGAGGAG
 [G,A]
 AGGAGCTAGGGCATGTTCTTCTGAGACTTGAGCGAGAGGAACCTTGGGAGTGGGAGGTTG
 TGGGGAAGTTAGAGGCTGCAAGGGCTGTTGAGGTAGTGAGAGGGACGGATCCCATGAGGA
 GTCTGGCATGGGGGCTCTGATTTAGCCTCTTCCCTGCAGTGGACAATGAGGAGGCTGCAC
 TGCTGCATGAAGAGGCTACCATGACTATTGAAGAGCTGCTGACACGCTACGGGCAGAACT
 GTCACAAGGGCCCTCCCCACAGCAAATCTGGAGGTGGGACAGGCGAGGAACCAGGGTCCC

GENEWISE ALIGNMENT:

gi 4505999 ref	1	MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPEL	
		MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSM	
		MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSM	
genomic	2134	aggtctccaagattggggggccccctgttgacgtcgtaggaacgcagt	
		tgcatcacactagcgagtgccgtctcagtcctaggtctatggagctgg	
		gtccccgcccggcgccgcgggccccccgacgcccggggggggcacg	
gi 4505999 ref	50	DSETAMFSVYDGHGG	EEVALYCA
		++ + G	EEVALYCA
		PLRGGNLTEKESGGW	EEVALYCA
genomic	2281	cccggacagagaggtGGTCCTCC	Intron 1 TAGGGggggttg
		ctgggatcaaaggg <1-----[2327 :25451]-1>	aatctagc
		gggctgtgggagcgag	gatcgctc
gi 4505999 ref	73	KYLPDI IKDQKAYKEGKLQK	ALEDAF
		KYLPDI IKDQKAYKEGKLQK	ALEDAF
		KYLPDI IKDQKAYKEGKLQK	ALEDAF
genomic	25478	atccgaaagcagtaggaccaGTCTGTC	Intron 2 CAGgtgggt
		aatcatataaacaagataa<0-----[25538:25851]-0>	taatcc
		attttccatggccgacgag	
gi 4505999 ref	99	LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD	
		LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD	
		LAI DAKLTTEEVIKELAQIAGRPTED EDEKEKVADEDD	
genomic	25870	tgaggataaggggaagcgaggccagggggagagggggg	
		tctacatccaattaatcatcgccaaaaaaatcaaaa	
		gttccagctaactaggagtagactgtataaaaattatt	
gi 4505999 ref	137		DNEEAALLHEEATMTIEELLTRY
			DNEEAALLHEEATMTIEELLTRY
		V:V[gtg]	DNEEAALLHEEATMTIEELLTRY
genomic	25984	GGTGAGTG	Intron 3 CAGTGgagggggcccgggaaaaggccact
		<1-----[25985:26642]-1>	aaaaccttaaacctctaattcga
			ctggtaggtagtcgttagggacc
gi 4505999 ref	161	GQNCHKGP PHSKSGGGTGEEPGSQGLNGEAGPEDSTRET PSEQENGPTAK	
		GQNCHKGP PHSKSGGGTGEEPGSQGLNGEAGPEDSTRET PSEQENGPTAK	
		GQNCHKGP PHSKSGGGTGEEPGSQGLNGEAGPEDSTRET PSEQENGPTAK	
genomic	26714	gcatcagcccaatgggagggcgctcgagggggcggttaagactcgagcaga	
		gaagaagccagacggcggaacgcagtagacgcaaccgacccaaagccca	
		ggctcgctcccatatgacgaagcgccctggaatgcatgattaaatccacg	
gi 4505999 ref	210	AYTGFSNSN SERGTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE	
		AYTGFSNSN SERGTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE	
		AYTGFSNSN SERGTEAGQVGEPIPTGEAGPSCSSASDKLPRVAKSKFFE	
genomic	26861	gtagtttatgcgagggcgggcgacaggggcttttgtgacccggatattg	
		cacgtccacaggcacgatgacgtccgacgcccgaatcgtcacatta	
		ccactcccgatgtgacattgtctcttgtgtcctactcggtattgcgctg	
gi 4505999 ref	259	DSEDESDEAE EEEEEEDSE	ECSE EEDGY

FIGURE 3, page 15 of 17

		DSEDESDEAEEDSE	ECSEEDGY
		DSEDESDEAEEDSE	ECSEEDGY
genomic	27008	gaggggtggggggggagGTAAGGG Intron 4 CAGgtaggggt	
		agaaacaacaaaaaga<0-----[27059:27638]-0>aggaaaaga	
		ctgtgatgggaagactg	accgagtcc
gi 4505999 ref	285	SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE	
		SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE	
		SSEEAENEDEDDTEEAEDDEEEEMMVPMEGKEE	
genomic	27666	aaggggaggggggagggggggggggaagcgaggagg	
		ggaacaaaaaaacaacaaaaaaatttcgtagaaa	
		ctggagtgatgtccggtagctaaaaagggaggacagg	
gi 4505999 ref	323	PGSDSGTTAVVALIRGKQLIVANAGD	
		PGSDSGTTAVVALIRGKQLIVANAGD	
		PGSDSGTTAVVALIRGKQLIVANAGD	
genomic	27780	GTGTGTG Intron 5 AAGcgtgagaaggggcacgactaggagg	
		<0-----[27780:28130]-0>cgaggcccttcttgaatttcacga	
		tctcttaagggcgaaggggtaccaac	
gi 4505999 ref	349	SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS	
		SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS	
		SRCVVSEAGKALDMSYDHKPEDEVELARIKNAGGKVTMDGRVNGGLNLS	
genomic	28209	tctggtgggagtgattgcacggggcgcaagggagaagcgaggcact	
		cggttcacgactatcaaaacaatatcgtaacggatctaggtaggtatc	
		tctgatgtcatagccttcaagtaaaaaccgtttcgccgtgaccgcccc	
gi 4505999 ref	398	RAI	DHFKRNKNLPPEEQMISAL
		RAI	DHFKRNKNLPPEEQMISAL
		RAI G:G[ggg]	DHFKRNKNLPPEEQMISAL
genomic	28356	agaGGTAAGGG Intron 6 CAGGGgcttaaaaacccggcaatgc	
		gct <1-----[28366:29164]-1> aataagaaatccaaattcct	
		act	ccctgacgcgatgaggtact
gi 4505999 ref	422	PDIKVLTLTDDHEFMVIACDGI	N
		PDIKVLTLTDDHEFMVIACDGI	N
		PDIKVLTLTDDHEFMVIACDGI W:W[tgg]	N
genomic	29227	cgaagcacaggcgtagagtggaTGGTGAGCA Intron 7 CAGGa	
		catatttctcaaaattttcgagt <2-----[29295:29541]-2> a	
		tccgggtctcctacgctcttc	t
gi 4505999 ref	446	VMSSQEVVDFIQSKISQRDENGELRLLSSIVEEL	
		VMSSQEVVDFIQSKISQRDENGELRLLSSIVEE+	
		VMSSQEVVDFIQSKISQRDENGELRLLSSIVEEV	
genomic	29546	gaaacgggggtactaaaccggaggccttttagggg	
		ttggaattattacatgagaaagatgttccttaat	
		ggccgatatctaagccgttatggtgagactgagg	
cDNA	1	MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEDAHNCIPEL	
		MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSME	
		MGAYLSQPNTVKCSGDGVGAPRLPLPYGFSAMQGWRVSMEEVRRQGP	
genomic	2134	aggtctccaagattggggggccccctgttgacgtcgttaggaacgcagt	
		tgcatacactagcgagtgccgtctcagtccttaggtctatgagctgg	
		gtccccgcccggcgcccgccggcccccgacgccccggggggggcacg	
cDNA	50	DSETAMFSVYDGHGG	EEVALYCA
		++ + G	EEVALYCA
		PLRGGNLTEKESGGW ~:G[ggg]	EEVALYCA
genomic	2281	cccggacagagaggtGGTCCTCC Intron 1 TAGGGgggggttg	
		ctgggatcaaaagggg <1-----[2327 :25451]-1> aatctagc	
		gggcgtgggagcgag	gacgctc

FIGURE 3, page 16 of 17

cDNA	73	KYLPDIKDQKAYKEGKLQK		ALEDAF
		KYLPDIKDQKAYKEGKLQK		ALEDAF
		KYLPDIKDQKAYKEGKLQK		ALEDAF
genomic	25478	atccgaaagcagtaggaccaGTCTGTC	Intron 2	CAGgtgggt
		aatcattaaaacaaagataa<0-----[25538:25851]-0>		ctaaact
		atctccatggccgacgagg		taatcc
cDNA	99	LAIDAKLTTEEVIKELAQIAGRPTEDDEKEKVADEDD		
		LAIDAKLTTEEVIKELAQIAGRPTEDDEKEKVADEDD		
		LAIDAKLTTEEVIKELAQIAGRPTEDDEKEKVADEDD		
genomic	25870	tgaggataaggggaagcgaggccaggggagagggggg		
		tctacatccaattaatcatcgccaaaaaaatcaaaa		
		gttccagctaactaggagtagactgtataaaaattatt		
cDNA	137			DHFKRNKNLPPEEQMISALPDI
				DHFKRNKNLPPEEQMISALPDI
		G:G[ggg]		DHFKRNKNLPPEEQMISALPDI
genomic	25984	GGTGAGTG	Intron 3	CAGGGgcttaaaaacccggcaatgccga
		<1-----[25985:29164]-1>		aataagaaatccaaattcctcat
				ccctgacgcgatgaggtacttcc
cDNA	161	KVLTLTDDHEFMVIACDGI		NVMS
		KVLTLTDDHEFMVIACDGI		NVMS
		KVLTLTDDHEFMVIACDGI		NVMS
genomic	29236	agcacaggcgtagagtggatGGTGAGCA	Intron 4	CAGGagaa
		attctcaaaattttcgagt	<2-----[29295:29541]-2>	attg
		gggtctcctacgctcttcc		tggc
cDNA	185	SQEVVDFIQSKISQDENGELRLLSSIVEEL		
		SQEVVDFIQSKISQDENGELRLLSSIVEE+		
		SQEVVDFIQSKISQDENGELRLLSSIVEEV		
genomic	29555	acgggggtactaaaccggaggccttttagggg		
		gaattattacatgagaaagatgttccttaat		
		cgatatctaagccgttatggtgagactgagg		